Partially Symmetric Functions are Efficiently Isomorphism-Testable

Eric Blais[*] Amit Weinstein[†] Yuichi Yoshida[‡]

December 30, 2011

Abstract

Given a function $f : \{0,1\}^n \to \{0,1\}$, the *$f$-isomorphism testing* problem requires a randomized algorithm to distinguish functions that are identical to f up to relabeling of the input variables from functions that are far from being so. An important open question in property testing is to determine for which functions f we can test f-isomorphism with a constant number of queries. Despite much recent attention to this question, essentially only two classes of functions were known to be efficiently isomorphism testable: symmetric functions and juntas.

We unify and extend these results by showing that all *partially symmetric* functions—functions invariant to the reordering of all but a constant number of their variables—are efficiently isomorphism-testable. This class of functions, first introduced by Shannon, includes symmetric functions, juntas, and many other functions as well. We conjecture that these functions are essentially the only functions efficiently isomorphism-testable.

To prove our main result, we also show that partial symmetry is efficiently testable. In turn, to prove this result we had to revisit the junta testing problem. We provide a new proof of correctness of the nearly-optimal junta tester. Our new proof replaces the Fourier machinery of the original proof with a purely combinatorial argument that exploits the connection between sets of variables with low influence and intersecting families.

Another important ingredient in our proofs is a new notion of *symmetric influence*. We use this measure of influence to prove that partial symmetry is efficiently testable and also to construct an efficient sample extractor for partially symmetric functions. We then combine the sample extractor with the testing-by-implicit-learning approach to complete the proof that partially symmetric functions are efficiently isomorphism-testable.

[*] School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA. Email: eblais@cs.cmu.edu

[†] Blavatnik School of Computer Science, Tel Aviv University, Tel Aviv, Israel. Email: amitw@tau.ac.il. Research supported in part by an ERC Advanced grant.

[‡] School of Informatics, Kyoto University and Preferred Infrastructure, Inc., Kyoto, Japan. Email: yyoshida@kuis.kyoto-u.ac.jp

1 Introduction

Property testing considers the following general problem: given a property $\mathcal{P}$, identify the minimum number of queries required to determine with high probability whether an input has the property $\mathcal{P}$ or whether it is far from $\mathcal{P}$. This question was first formalized by Rubinfeld and Sudan [27].

Definition 1 ([27])**.** Let $\mathcal{P}$ be a set of Boolean functions. An *ϵ-tester* for $\mathcal{P}$ is a randomized algorithm which queries an unknown function $f : \{0,1\}^n \to \{0,1\}$ on a small number of inputs and

(i) Accepts with probability at least 2/3 when $f \in \mathcal{P}$;

(ii) Rejects with probability at least 2/3 when f is ϵ-far from $\mathcal{P}$,

where f is *ϵ-far* from $\mathcal{P}$ if $\mathrm{dist}(f,g) := |\{x \in \{0,1\}^n \mid f(x) \neq g(x)\}| \geq \epsilon 2^n$ holds for every $g \in \mathcal{P}$.

Goldreich, Goldwasser, and Ron [22] extended the scope of this definition to graphs and other combinatorial objects. Since then, the field of property testing has been very active. For an overview of recent developments, we refer the reader to the surveys [25, 26] and the book [21].

A notable achievement in the field of property testing is the complete characterization of graph properties that are testable with a constant number of queries [5]. An ambitious open problem is obtaining a similar characterization for properties of Boolean functions. Recently there has been a lot of progress on the restriction of this question to properties that are closed under linear or affine transformations [6, 23]. More generally, one might hope to settle this open problem for all properties of Boolean functions that are closed under relabeling of the input variables.

An important sub-problem of this open question is function isomorphism testing. Given a Boolean function f, the *f-isomorphism testing* problem is to determine whether a function g is isomorphic to f—that is, whether it is the same up to relabeling of the input variables—or *far* from being so. A natural goal, and the focus of this paper, is to characterize the set of functions for which isomorphism testing can be done with a constant number of queries.

Previous work. The function isomorphism testing problem was first raised by Fischer et al. [17]. They observed that fully symmetric functions are trivially isomorphism testable with a constant number of queries. They also showed that every *k-junta*, that is every function which depends on at most k of the input variables, is isomorphism testable with $\mathrm{poly}(k)$ queries. This bound was recently improved by Chakraborty et al. [12], who showed that $O(k \log k)$ suffice. In particular, these results imply that juntas on a constant number of variables are isomorphism testable with a constant number of queries.

The first lower bound for isomorphism testing was also provided by Fischer et al. [17]. They showed that for small enough values of k, testing isomorphism to a k-linear function (i.e., a function that returns the parity of k variables) requires $\Omega(\log k)$ queries.[1] Following a series of recent works [20, 8, 9], the exact query complexity for testing isomorphism to k-linear functions has been determined to be $\tilde{\Theta}(\min(k, n-k))$.

More general lower bounds for isomorphism testing were obtained by O'Donnell and the first author [10]. In particular, they showed that testing isomorphism to *any* k-junta that is *far* from being a $(k-1)$-junta requires $\Omega(\log \log k)$ queries. This lower bound gives a large family of functions for which testing isomorphism requires a super-constant number of queries. Alon et al. have shown that in fact the query complexity of testing isomorphism is $\tilde{\Theta}(n)$ for almost every function [4] (see also [3, 12]).

[1] More precisely, they showed that non-adaptive testers require $\tilde{\Omega}(\sqrt{k})$ queries. Here and in the rest of this section, tilde notation is used to hide logarithmic factors.

Partially symmetric functions. As seen above, the only functions which we know are isomorphism testable with a constant number of queries are fully symmetric functions and juntas. Our motivation for the current work was to see if we can unify and generalize the results to encompass a larger class of functions. While symmetric functions and juntas may seem unrelated, there is in fact a strong connection. Symmetric functions, of course, are invariant under any relabeling of the input variables. Juntas satisfy a similar but slightly weaker invariance property. For every k-junta, there is a set of at least $n-k$ variables such that the function is invariant to any relabeling of these variables. Functions that satisfy this condition are called *partially symmetric.*

Definition 2 (Partially symmetric functions)**.** For a subset $J \subseteq [n] := \{1, \ldots, n\}$, a function $f : \{0,1\}^n \to \{0,1\}$ is *J-symmetric* if permuting the labels of the variables of J does not change the function. Moreover, f is called *t-symmetric* if there exists $J \subseteq [n]$ of size at least t such that f is J-symmetric.

Shannon first introduced partially symmetric functions as part of his investigation on the circuit complexity of Boolean functions [28]. He showed that while most functions require an exponential number of gates to compute, every partially symmetric function can be implemented much more efficiently. Research on the role of partial symmetry in the complexity of implementing functions in circuits, binary decision diagrams, and other models has remained active ever since [13, 24]. Our results suggest that studying partially symmetric functions may also yield greater understanding of property testing on Boolean functions.

Our results. The set of partially symmetric functions includes both juntas and symmetric functions, but the set also contains many other functions as well. A natural question is whether this entire class of functions is isomorphism testable with a constant number of queries. Our first main result gives an affirmative answer to this question.

Theorem 1. *For every $(n-k)$-symmetric function $f : \{0,1\}^n \to \{0,1\}$ there exists an ϵ-tester for f-isomorphism that performs $O(k \log k/\epsilon^2)$ queries.*

A simple modification of an argument in Alon et al. [4] can be used to show that the bound in the above theorem is tight up to logarithmic factors. Indeed by this argument, testing isomorphism to almost every $(n-k)$-symmetric function requires $\Omega(k)$ queries.

We believe that the theorem might also be best possible in a different way. That is, we conjecture that the set of partially symmetric functions is essentially the set of functions for which testing isomorphism can be done with a constant number of queries. We discuss this conjecture with some supporting evidence in Section 6.

The proof of our first main theorem follows the general outline of the proof that isomorphism testing to juntas can be done in a constant number of queries. The observation which allows us to make this connection is the fact that partially symmetric functions can be viewed as junta-like functions. More precisely, an $(n-k)$-symmetric function is a function that has k special variables where for each assignment for these variables, the restricted function is fully symmetric on the remaining $n-k$ variables.

The proof for testing isomorphism of juntas has two main components. The first is an efficient junta testing algorithm. This enables us to reject functions that are far from being juntas. The second is a query efficient sampler of the "core" of the input function given that the function is close to a junta. The sampler can then be used in order to verify if the two juntas are indeed isomorphic. We generalize both of these components for partially symmetric functions.

Our second main result, and the first component of the isomorphism tester, is an efficient algorithm for testing partial symmetry.

Theorem 2. *The property of being $(n-k)$-symmetric for $k < n/10$ is testable with $O(\frac{k}{\epsilon}\log\frac{k}{\epsilon})$ queries.*

The natural approach for proving this theorem is to try generalize the result on junta testing in [7]. That result heavily relied on the notion of influence of variables. The *influence* of a set S of variables in a function f is the probability that $f(x) \neq f(y)$ when x is chosen uniformly at random and y is obtained from x by re-randomizing the values of x_i for each $i \in S$. The notion of influence characterizes juntas: when f is a k-junta, there is a set of size $n-k$ whose influence is 0, whereas when f is ϵ-far from being a k-junta, every set of size $n-k$ has influence at least ϵ.

We introduce a different notion of influence which we call *symmetric influence*. The symmetric influence of a set S of variables in f is the probability that $f(x) \neq f(y)$ when x is chosen uniformly at random and y is obtained from x by permuting the values of $\{x_i\}_{i \in S}$. This notion characterizes partially symmetric functions and satisfies several other useful properties. We provide the details in Section 3.

The proof of junta testing also relies on nice properties of the Fourier representation of the notion of influence. While symmetric influence has a clean Fourier representation, unfortunately it does not have the properties needed to carry over the proof in [7] to the setting of partially symmetric functions. Instead, we must come up with a new proof technique.

Our proof of Theorem 2 uses a new connection to intersecting families. A family $\mathcal{F}$ of subsets of $[n]$ is *t-intersecting* if for every pair of sets $S, T \in \mathcal{F}$, their intersection size is at least $|S \cap T| \geq t$. This notion was introduced by Erdős, Ko, and Rado and a sequence of works led to the complete characterization of the maximum size of t-intersecting families that contain sets of fixed size [16, 18, 30, 2]. Dinur, Safra, and Friedgut recently extended those results to give bounds on the biased measure of intersecting families [15, 19].

Using results in intersecting families, we obtain a new and improved proof for the main lemma at the heart of the junta testing result [7]. We describe the new proof and the connection to intersecting families in Section 2. Most importantly, the same technique can also be extended to complete the proof of Theorem 2. We present this proof in Section 4.

The second and final component of the isomorphism test for partially symmetric functions is an efficient way to sample the core of such functions. An $(n-k)$-symmetric function f, which is symmetric over a set $J \subseteq [n]$ of size $|J| = n-k$, has a concise representation as a function $f_{\text{core}} : \{0,1\}^k \times \{0, 1, \ldots, n-k\} \to \{0,1\}$ which we call the *core* of f. The core is the restriction of f to the variables in $\overline{J}$ (in the natural order), with the additional Hamming weight of the variables in J. To determine if two partially symmetric functions are isomorphic, it suffices to determine whether their cores are isomorphic. We do so with the help of an efficient sample extractor.

Definition 3. A (1 query) *δ-sampler* for the $(n-k)$-symmetric function $f : \{0,1\}^n \to \{0,1\}$ is a randomized algorithm that queries f on a single input and returns a triplet $(x, w, z) \in \{0,1\}^k \times \{0, 1, \ldots, n-k\} \times \{0,1\}$ where

- The distribution of (x, w) is δ-close, in total variation distance, to x being uniform over $\{0,1\}^k$ and w being binomial over $\{0, 1, \ldots, n-k\}$ independently, and
- $z = f_{\text{core}}(x, w)$ with probability at least $1 - \delta$.

Our third main result is that for any $(n-k)$-symmetric function f, there is a query-efficient algorithm for constructing a δ-sampler for f.

Theorem 3. *Let $f : \{0,1\}^n \to \{0,1\}$ be $(n-k)$-symmetric with $k < n/10$. There is an algorithm that queries f on $O(\frac{k}{\eta\delta}\log\frac{k}{\eta\delta})$ inputs and with probability at least $1-\eta$ outputs a δ-sampler for f.*

This theorem is a generalization of a recent result of Chakraborty et al. [11], who gave a similar construction for sampling the core of juntas. Their result has many applications related to testing by implicit learning [14]. Our result may be of independent interest for similar such applications. We elaborate on this topic and present the proof of Theorem 3 in Section 5.

2 Intersecting families and testing juntas

We begin by revisiting the problem of junta testing. In this section, we give a new proof of the correctness of the k-junta tester first introduced in [7]. At a high level, the junta tester is quite simple: it partitions the set of indices into a large enough number of parts, then tries to identify all the parts that contain a relevant variable. If at most k such parts are found, the test accepts; otherwise it rejects. The algorithm is described in JUNTA-TEST. (See [7] for more details.)

Algorithm 1 JUNTA-TEST(f, k, ϵ)

```
1: Create a random partition $\mathcal{I}$ of the set $[n]$ into $r = \Theta(k^2)$ parts, and initialize $J = \emptyset$.
2: for each $i = 1$ to $\Theta(k/\epsilon)$ do
3:    Sample $x, y \in \{0,1\}^n$ uniformly at random.
4:    if $f(x) \neq f(x_J y_{\overline{J}})$ then
5:       Use binary search to find a set $I \in \mathcal{I}$ that contains a relevant variable.
6:       Set $J := J \cup I$.
7:       if $J$ is the union of $> k$ parts then reject.
8: Accept.
```

It is clear that the JUNTA-TEST always accepts k-juntas. The non-trivial part of the analysis involves showing that functions that are far from k-juntas are rejected by the tester with sufficiently high probability. To do so, we must argue that the inequality in Step 4 is satisfied with non-negligible probability whenever f is far from k-juntas and J is the union of at most k parts. This is accomplished by considering the influence of variables in a function.

The *influence* of the set $J \subseteq [n]$ of variables in the function $f : \{0,1\}^n \to \{0,1\}$ is

$$\mathrm{Inf}_f(J) = \Pr_{x,y}[f(x) \neq f(x_{\overline{J}} y_J)] \ ,$$

where $x_{\overline{J}} y_J$ is the vector $z \in \{0,1\}^n$ obtained by setting $z_i = y_i$ for every $i \in J$ and $z_i = x_i$ for every $i \in [n] \setminus J$. By definition, the probability that the inequality in Step 4 is satisfied is exactly $\mathrm{Inf}_f(\overline{J})$. To complete the analysis of correctness of the algorithm, we want to show that when f is ϵ-far from k-juntas with high probability over the choice of the random partition $\mathcal{I}$, if J is the union of at most k parts in $\mathcal{I}$, then $\mathrm{Inf}(\overline{J}) \geq \frac{\epsilon}{4}$. We do so by exploiting only a couple basic facts about the notion of influence.

Lemma 1 (Fischer et al. [17]). *For every $f : \{0,1\}^n \to \{0,1\}$ and every $J, K \subseteq [n]$,*

$$\mathrm{Inf}_f(J) \leq \mathrm{Inf}_f(J \cup K) \leq \mathrm{Inf}_f(J) + \mathrm{Inf}_f(K) \ .$$

Furthermore, if f is ϵ-far from k-juntas and $|J| \leq k$, then $\mathrm{Inf}_f(\overline{J}) \geq \epsilon$.

We also use the fact that the family of sets $J \subseteq [n]$ whose complements have small influence form an intersecting family. For a fixed $t \geq 1$, a family $\mathcal{F}$ of subsets of $[n]$ is called *t-intersecting* if any two sets J and K in $\mathcal{F}$ have intersection size $|J \cap K| \geq t$. Much of the work in this area focused on bounding the size of t-intersecting families that contain only sets of a fixed size. Dinur and Safra [15] considered general families and asked what the maximum *p-biased measure* of such families can be. For $0 < p < 1$, this measure is defined as $\mu_p(\mathcal{F}) := \Pr_J[J \in \mathcal{F}]$ where the probability over J is obtained by including each coordinate $i \in [n]$ in J independently with probability p. They showed that 2-intersecting families have small p-biased measure [15] and Friedgut showed how the same result also extends to t-intersecting families for $t > 2$ [19].

Theorem 4 (Dinur and Safra [15]; Friedgut [19])**.** *Let $\mathcal{F}$ be a t-intersecting family of subsets of $[n]$ for some $t \geq 1$. For any $p < \frac{1}{t+1}$, the p-biased measure of $\mathcal{F}$ is bounded by $\mu_p(\mathcal{F}) \leq p^t$.*

We are now ready to complete the analysis of Junta-Test.

Lemma 2. *Let $f : \{0,1\}^n \to \{0,1\}$ be a function ϵ-far from k-juntas and $\mathcal{I}$ be a random partition of $[n]$ into $r = c \cdot k^2$ parts, for some large enough constant c. Then with probability at least $5/6$, $\mathrm{Inf}_f(\overline{J}) \geq \epsilon/4$ for any union J of k parts from $\mathcal{I}$.*

Proof. For $0 \leq t \leq \frac{1}{2}$, let $\mathcal{F}_t = \{J \subseteq [n] : \mathrm{Inf}_f(\overline{J}) < t\epsilon\}$ be the family of all sets whose complements have influence at most $t\epsilon$. For any two sets $J, K \in \mathcal{F}_{1/2}$, the sub-additivity of influence implies that

$$\mathrm{Inf}_f(\overline{J \cap K}) = \mathrm{Inf}_f(\overline{J} \cup \overline{K}) \leq \mathrm{Inf}_f(\overline{J}) + \mathrm{Inf}_f(\overline{K}) < 2 \cdot \tfrac{1}{2}\epsilon = \epsilon \ .$$

But f is ϵ-far from k-juntas, so every set $S \subseteq [n]$ of size $|S| \leq k$ satisfies $\mathrm{Inf}_f(\overline{S}) \geq \epsilon$. Therefore, $|J \cap K| > k$ and, since this argument applies to every pair of sets in the family, $\mathcal{F}_{1/2}$ is a $(k+1)$-intersecting family.

Let us now consider two separate cases: when $\mathcal{F}_{1/2}$ contains a set of size less than $2k$; and when it does not. In the first case, let $J \in \mathcal{F}_{1/2}$ be one of the sets of size $|J| < 2k$. With high probability, the set J is completely separated by the partition $\mathcal{I}$. When this event occurs, then for every other set $K \in \mathcal{F}_{1/2}$, $|J \cap K| \geq k+1$, which means that K is not covered by any union of k parts in $\mathcal{I}$. Therefore, with high probability f is $\frac{\epsilon}{2}$-far (and thus also $\frac{\epsilon}{4}$-far) from k-part juntas with respect to $\mathcal{I}$, as we wanted to show.

Consider now the case where $\mathcal{F}_{1/2}$ contains only sets of size at least $2k$. Then we claim that $\mathcal{F}_{1/4}$ is a $2k$-intersecting family: otherwise, we could find sets $J, K \in \mathcal{F}_{1/4}$ such that $|J \cap K| < 2k$ and $\mathrm{Inf}_f(\overline{J \cap K}) \leq \mathrm{Inf}_f(\overline{J}) + \mathrm{Inf}_f(\overline{K}) < \frac{\epsilon}{2}$, contradicting our assumption.

Let $J \subseteq [n]$ be the union of k parts in $\mathcal{I}$. Since $\mathcal{I}$ is a random partition, J is a random subset obtained by including each element of $[n]$ in J independently with probability $p = \frac{k}{r} < \frac{1}{2k+1}$. By Theorem 4,

$$\Pr_{\mathcal{I}}[\mathrm{Inf}_f(\overline{J}) < \tfrac{\epsilon}{4}] = \Pr[J \in \mathcal{F}_{1/4}] = \mu_{k/r}(\mathcal{F}_{1/4}) \leq (k/r)^{2k} \ .$$

Applying the union bound over the possible choices of J, we get that f is $\frac{\epsilon}{4}$-close to a k-part junta with respect to $\mathcal{I}$ with probability at most

$$\binom{r}{k}\left(\frac{k}{r}\right)^{2k} \leq \left(\frac{er}{k}\right)^k \left(\frac{k}{r}\right)^{2k} \leq \left(\frac{ek}{r}\right)^k = O(k^{-k}) \ .$$

□

3 Symmetric influence

The main focus of this paper is partially symmetric functions, that is, functions invariant under any reordering of the variables of some set $J \subseteq [n]$. Let $\mathcal{S}_J$ denote the set of permutations of $[n]$ which only move elements from the set J. A function $f : \{0,1\}^n \to \{0,1\}$ is J-symmetric if $f(x) = f(\pi x)$ for every input x and a permutation $\pi \in \mathcal{S}_J$, where πx is the vector whose $\pi(i)$-th coordinate is x_i.

For better analyzing partially symmetric functions, we introduce a new measure named *symmetric influence.* The symmetric influence of a set measures how invariant the function is to reordering of the elements in that set.

Definition 4 (Symmetric influence)**.** The *symmetric influence* of a set $J \subseteq [n]$ of variables in a Boolean function $f : \{0,1\}^n \to \{0,1\}$ is defined as

$$\mathrm{SymInf}_f(J) = \Pr_{x \in \{0,1\}^n, \pi \in \mathcal{S}_J}[f(x) \neq f(\pi x)] \ .$$

It is not hard to see that in fact a function f is t-symmetric iff there exists a set J of size t such that $\mathrm{SymInf}_f(J) = 0$. A much stronger connection, however, exists between these properties as we will shortly describe.

Before showing some nice properties of symmetric influence, we mention that it also has a simple representation using Fourier coefficients of the function. Although we do not use the representation in this paper, we feel it might be of independent interest. See Appendix A.1 for details.

Lemma 3. *Given a function $f : \{0,1\}^n \to \{0,1\}$ and a subset $J \subseteq [n]$, let f_J be the J-symmetric function closest to f. Then, the symmetric influence of J satisfies*

$$\mathrm{dist}(f, f_J) \leq \mathrm{SymInf}_f(J) \leq 2 \cdot \mathrm{dist}(f, f_J) \ .$$

Proof. For every weight $0 \leq w \leq n$ and $z \in \{0,1\}^{|\overline{J}|}$, define the layer $L^w_{\overline{J}\leftarrow z} := \{x \in \{0,1\}^n \mid |x| = w \wedge x_{\overline{J}} = z\}$ to be the vectors of Hamming weight w which identify with z over the set $\overline{J}$ (where $|L^w_{\overline{J}\leftarrow z}| = \binom{|J|}{w-|z|}$ if $|z| \leq w \leq |J| + |z|$ or 0 otherwise). Let $p^w_z \in [0, \frac{1}{2}]$ be the fraction of the vectors in $L^w_{\overline{J}\leftarrow z}$ one has to modify in order to make the restriction of f over $L^w_{\overline{J}\leftarrow z}$ constant.

With these notations, we can restate the definition of the symmetric influence of J as follows.

$$\begin{aligned}\mathrm{SymInf}_f(J) &= \sum_z \sum_w \Pr_{x \in \{0,1\}^n}[x \in L^w_{\overline{J}\leftarrow z}] \cdot \Pr_{x \in \{0,1\}^n, \pi \in \mathcal{S}_J}[f(x) \neq f(\pi x) \mid x \in L^w_{\overline{J}\leftarrow z}] \\ &= \frac{1}{2^n} \sum_z \sum_w |L^w_{\overline{J}\leftarrow z}| \cdot 2p^w_z(1 - p^w_z) \ .\end{aligned}$$

This holds as in each such layer, the probability that x and πx would result in two different outcomes is the probability that x would be chosen out of the smaller part and πx from the complement, or vise versa.

The function f_J can be obtained by modifying f at p^w_z fraction of the inputs in each layer $L^w_{\overline{J}\leftarrow z}$, as each layer can be addressed separately and we want to modify as few inputs as possible. By this observation, we have the following equality.

$$\mathrm{dist}(f, f_J) = \frac{1}{2^n} \sum_z \sum_w |L^w_{\overline{J}\leftarrow z}| \cdot p^w_z \ .$$

But since $1 - p^w_z \in [\frac{1}{2}, 1]$, we have that $p^w_z \leq 2p^w_z(1 - p^w_z) \leq 2p^w_z$ and therefore $\mathrm{dist}(f, f_J) \leq \mathrm{SymInf}_f(J) \leq 2 \cdot \mathrm{dist}(f, f_J)$ as required. □

Corollary 1. *Let $f : \{0,1\}^n \to \{0,1\}$ be a function that is ϵ-far from being t-symmetric. Then for every set $J \subseteq [n]$ of size $|J| \geq t$, $\mathrm{SymInf}_f(J) \geq \epsilon$ holds.*

Proof. Fix $J \subseteq [n]$ of size $|J| \geq t$ and let g be a J-symmetric function closest to f. Since g is symmetric on any subset of J, it is in particular t-symmetric and therefore $\mathrm{dist}(f, g) \geq \epsilon$ as f is ϵ-far from being t-symmetric. Thus, by Lemma 3, $\mathrm{SymInf}_f(J) \geq \mathrm{dist}(f, g) \geq \epsilon$ holds. □

Corollary 1 demonstrates the strong connection between symmetric influence and the distance from being partially symmetric, similar to the second part of Lemma 1 for influence and juntas. The additional properties of influence used in Section 2 are monotonicity and sub-additivity (Lemma 1). The following lemmas show that the same properties (approximately) hold for symmetric influence. See Appendices A.2 and A.3 for the proofs of both lemmas.

Lemma 4 (Monotonicity)**.** *For any function $f : \{0,1\}^n \to \{0,1\}$ and any sets $J \subseteq K \subseteq [n]$,*

$$\mathrm{SymInf}_f(J) \leq \mathrm{SymInf}_f(K) \ .$$

Lemma 5 (Weak sub-additivity)**.** *There is a universal constant c such that, for any constant $0 < \gamma < 1$, a function $f : \{0,1\}^n \to \{0,1\}$, and sets $J, K \subseteq [n]$ of size at least $(1-\gamma)n$,*

$$\mathrm{SymInf}_f(J \cup K) \leq \mathrm{SymInf}_f(J) + \mathrm{SymInf}_f(K) + c\sqrt{\gamma} \ .$$

4 Testing partial symmetry

Let us now return to the problem of testing partial symmetry. The goal of this section is to introduce an efficient tester for this property by combining the ideas from Sections 2 and 3.

We begin by introducing the testing algorithm Partially-Symmetric-Test. This algorithm is conceptually very similar to the junta tester in Section 2. Again, the main idea is to partition the variables into $O(k^2)$ parts and identify the parts that contain "asymmetric" variables. More precisely, given a function $f : \{0,1\}^n \to \{0,1\}$, let us write $\mathrm{core}(f) \subseteq [n]$ to be the maximum set J of variables such that f is J-symmetric. We call the variables in $\mathrm{core}(f)$ *symmetric* and the variables in $[n] \setminus \mathrm{core}(f)$ are called *asymmetric*. The function is $(n-k)$-symmetric iff it contains at most k asymmetric variables. The algorithm exploits this characterization by trying to identify $k+1$ parts that contain asymmetric variables.

Algorithm 2 Partially-Symmetric-Test(f, k, ϵ)

```
Create a random partition $\mathcal{I}$ of $[n]$ into $r = \Theta(k^2/\epsilon^2)$ parts, and initialize $J := \emptyset$.
Pick a random workspace $W \in \mathcal{I}$, and if $|W| < \frac{n}{2r}$ then fail.
for each $i = 1$ to $\Theta(k/\epsilon)$ do
    Let $I :=$ Find-Asymmetric-Set$(f, \mathcal{I}, J, W)$.
    if $I \neq \emptyset$ then
        Set $J := J \cup I$.
        if $J$ is the union of $> k$ parts then reject.
Accept.
```

There are two main differences in the analysis of Partially-Symmetric-Test and of Junta-Test in Section 2. The first is that we can no longer use a simple binary search algorithm to identify the parts that contain asymmetric variables, as we need to maintain the Hamming weight of our queries. To overcome this challenge, we introduce the Find-Asymmetric-Set function, which satisfies the following properties.

Lemma 6. *Let f be a function, $\mathcal{I}$ be a partition of $[n]$ into r parts, $W \in \mathcal{I}, |W| \geq \frac{n}{2r}$ be a workspace, and J be a union of parts from $\mathcal{I} \setminus \{W\}$. Then, there exists an algorithm* Find-Asymmetric-Set$(f, \mathcal{I}, J, W)$ *which performs $O(\log r)$ queries such that*

- *With probability $\mathrm{SymInf}_f(\overline{J})$, the algorithm returns a set $I \in \mathcal{I} \setminus \{W\}$ disjoint to J; otherwise it returns $\emptyset$.*

- *If W has no asymmetric variable and $I \in \mathcal{I}$ is returned, then I has an asymmetric variable.*

Due to space constraints, we provide a rough sketch of the algorithm and defer the details and analysis to Appendix B.1. Find-Asymmetric-Set generates a random pair of $x \in \{0,1\}^n$ and $\pi \in \mathcal{S}_{\overline{J}}$ and checks whether $f(x) \neq f(\pi x)$. When this occurs, which happens with probability at least ϵ when $\mathrm{SymInf}_f(\overline{J}) \geq \epsilon$, we know there exists some asymmetric variable in $\overline{J}$. In order to identify a part $I \in \mathcal{I}$, disjoint to J and the workspace W, which contains an asymmetric variable we iteratively change x to πx. In each step, we only *permute* bits in one part $I \in \mathcal{I}$ and the workspace W. Since $f(x) \neq f(\pi x)$, we can find using binary search a set I, disjoint to J, such that permuting bits in $I \cup W$ changes the value of f. By our assumption, W has no asymmetric variables and therefore I must contain such a variable.

The second and more important challenge in the analysis of Partially-Symmetric-Test is the use of symmetric influence (rather than influence). Similar to Lemma 2 for influence, we prove that if a function is far from being $(n-k)$-symmetric, then it is also far from being symmetric on any union of all but k parts of a random partition (assuming it has enough parts). The formal statement is given in Lemma 7, where its proof follows a very similar technique to that of Lemma 2.

Lemma 7. *Let $f : \{0,1\}^n \to \{0,1\}$ be a function ϵ-far from $(n-k)$-symmetric and $\mathcal{I}$ be a random partition of $[n]$ into $r = c \cdot k^2/\epsilon^2$ parts, for some large enough constant c. Then with probability at least $8/9$, $\mathrm{SymInf}_f(\overline{J}) \geq \frac{\epsilon}{9}$ holds for any union J of k parts.*

The main difference between this proof and the one of Lemma 2 arises from the weak sub-additivity of symmetric influence. In light of this difference, our definition of families of sets whose complement has small symmetric influence includes only sets which are not too big. We use the observation that adding sets which contain elements of a family does not change its existing intersection. In addition, due to the additive factor of the sub-additivity we prove a slightly weaker result where the symmetric influence is at least $\epsilon/9$ and not $\epsilon/4$. The complete proof of Lemma 7 is deferred to Appendix B.2.

We can now complete the proof that partial symmetry is efficiently testable.

Proof of Theorem 2. Note that $|W| \geq \frac{n}{2r}$ indeed holds with probability at least 8/9 from Chernoff bound. By Lemma 6, Find-Asymmetric-Set performs $O(\log \frac{k}{\epsilon})$ queries according to our choice of r, and therefore the query complexity of Partially-Symmetric-Test is $O(\frac{k}{\epsilon} \log \frac{k}{\epsilon})$.

Suppose f is an $(n-k)$-symmetric function. The probability that W contains an asymmetric variable is at most $k/r \leq 2/9$. Conditioned this did not occur, every set returned by Find-Asymmetric-Set contains an asymmetric variable. Since there are at most k such variables, J would be the union of at most k sets and we would accept.

Suppose f is a function ϵ-far from being $(n-k)$-symmetric. From Lemma 7, with probability at least 8/9, $\mathrm{SymInf}_f(\overline{J}) \geq \epsilon/9$ holds while J consists of at most k parts. Conditioned on that, by executing Find-Asymmetric-Set $O(k/\epsilon)$ times we obtain more than k parts with probability at least 8/9, according to Lemma 6. Thus, we reject with probability at least 2/3. □

5 Isomorphism testing of partially symmetric functions

In this section we prove that isomorphism testing of partially symmetric functions can be done with a constant number of queries. The algorithm we describe consists of two main components, and follow a similar approach to the one used in [12] when they showed juntas are isomorphism testable. The first, which we already described in Section 4, is an efficient tester for the property of being partially symmetric. Once we know the input function is indeed close to being partially symmetric, we can verify it is isomorphic (or at least very close) to the correct one. The second component of the algorithm is therefore an efficient sampler from the *core* of a function which is (close to) partially symmetric. Comparing the cores of two partially symmetric functions suffices to identify if two such functions are isomorphic or far from it.

Ideally, when sampling the core of a partially symmetric function f, we would like to sample it according to the marginal distribution of sampling f at a uniform input $x \in \{0,1\}^n$. We denote this marginal distribution over $\{0,1\}^k \times \{0,1,\dots,n-k\}$ by $\mathcal{D}^*_{k,n}$, which is in fact uniform over $\{0,1\}^k$ and binomial over $\{0,1,\dots,n-k\}$, independently.

In our scenario, sampling the core of a function according to this distribution is not possible since we do not know the exact location of all the k asymmetric variables. Instead, we use the knowledge discovered by the partial symmetry tester, i.e., sets with asymmetric variables. Given these sets, we are able to define a sampling distribution over $\{0,1\}^n$ such that we know the input of the core for each query, and whose marginal distribution over the core is close enough to $\mathcal{D}^*_{k,n}$.

Definition 5. Let $\mathcal{I}$ be some partition of $[n]$ into an odd number of parts and let $W \in \mathcal{I}$ be the workspace. Define the distribution $\mathcal{D}^W_{\mathcal{I}}$ over $\{0,1\}^n$ to be as follows. Pick a random Hamming weight w according to the binomial distribution over $\{0,\dots,n\}$ and output, if exists, a random $x \in \{0,1\}^n$ of Hamming weight $|x| = w$ such that for every $I \in \mathcal{I} \setminus \{W\}$, either $x_I \equiv 0$ or $x_I \equiv 1$. When no such x exists, return the all zeros vector.

The sampling distribution which we just defined, together with the random choice of the partition and workspace, satisfies the following two important properties. The first, being close to uniform over the inputs of the function. The second, having a marginal distribution over the core of a partially symmetric function close to $\mathcal{D}^*_{k,n}$. These properties are formally written here as Proposition 1, whose proof is rather technical and appears in Appendix C.1.

Proposition 1. *Let $J = \{j_1,\dots,j_k\} \subseteq [n]$ be a set of size k, and $r = \Omega(k^2)$ be odd. If $x \sim \mathcal{D}^W_{\mathcal{I}}$ for a random partition $\mathcal{I}$ of $[n]$ into r parts and a random workspace $W \in \mathcal{I}$, then*

- *x is $o(1/n)$-close to being uniform over $\{0,1\}^n$, and*
- *$(x_J, |x_{\overline{J}}|)$ is c/k-close to being distributed according to $\mathcal{D}^*_{k,n}$, for our choice of $0 < c < 1$.*

We are now ready to describe the algorithm for isomorphism testing of $(n-k)$-symmetric functions. Given an $(n-k)$-symmetric function f, the algorithm tests whether the input function g is isomorphic to f or ϵ-far from being so.

Algorithm 3 PARTIALLY-SYMMETRIC-ISOMORPHISM-TEST(f, k, g, ϵ)

```
Perform PARTIALLY-SYMMETRIC-TEST$(g, k, \epsilon/1000)$ and reject if failed.
Let $\mathcal{I}$ and $W \in \mathcal{I}$ be the partition and workspace used by the algorithm.
Let $J$ be the union of the $k$ parts identified by the algorithm.
for each $i = 1$ to $\Theta(k \log k/\epsilon^2)$ do
    Query $g(x)$ at a random $x \sim \mathcal{D}^W_{\mathcal{I}}$
Accept iff $(1-\epsilon/2)$-fraction of the queries are consistent with some isomorphism $f_\pi$ of $f$, which
    maps the asymmetric variables of $f$ into all $k$ parts of $J$.
```

We provide here a sketch of the analysis of the algorithm. See Appendix C.2 for the formal analysis and complete proof of Theorem 1. The first case to analyze is when g is rejected by PARTIALLY-SYMMETRIC-TEST, which implies that with good probability it is not $(n-k)$-symmetric and in particular not isomorphic to f. Assume now that PARTIALLY-SYMMETRIC-TEST did not reject and therefore g is likely to be $\epsilon/1000$-close to being $(n-k)$-partially symmetric. Let $\mathcal{I}, W$ and J be the partition, workspace and union of k parts identified by the algorithm. The main idea of the proof is showing that with good probability, there exists a function h that (a) is $\epsilon/250$-close to g, and (b) is $(n-k)$-symmetric with asymmetric variables contained in J and separated by $\mathcal{I}$. We prove the existence of this function h using the properties of symmetric influence presented in Section 4. Assuming such h exists, we use Proposition 1 in order to show that our queries to g, according to the sampling distribution, are in fact $\epsilon/10$-close to querying h's core.

We now consider the following two cases. If g is isomorphic to f, then for some isomorphism f_π of f, which maps the asymmetric variables of f into the parts of J, it holds that $\mathrm{dist}(f_\pi, h) \leq \mathrm{dist}(f_\pi, g) + \mathrm{dist}(g, h) \leq \epsilon/500 + \epsilon/250$. Notice that we cannot assume that $g = f_\pi$ as it is possible that one of the asymmetric variables of g are not in J (but the distance must be small). If g was ϵ-far from being isomorphic to f, then for every isomorphism f_π of f,

$$\mathrm{dist}(f_\pi, h) \geq \mathrm{dist}(f_\pi, g) - \mathrm{dist}(g, h) \geq \epsilon - \epsilon/250 \ .$$

Given that there are only $k!$ isomorphisms of f we need to consider, performing $\Theta(k \log k/\epsilon^2)$ queries suffices for returning the correct answer in both cases, with good probability.

As we outlined above, we in fact build an efficient sampler for the core of $(n-k)$-symmetric functions (or functions close to being so). Given the parts identified by PARTIALLY-SYMMETRIC-TEST, assuming it did not reject, we can sample the function's core by querying it at a single location, where the distribution over the core's inputs is close to $\mathcal{D}^*_{k,n}$. The algorithm and proof of Theorem 3 are deferred to Appendix C.3.

6 Discussion

We showed that every partially symmetric function is isomorphism testable with a constant number of queries. It's easy to see that functions that are "close" to partially symmetric can also be isomorphism-tested with a constant number of queries. We believe that our result not only unifies the previous classes of functions efficiently isomorphism-testable, but that it includes essentially *all* of these functions.

Conjecture 1. *Let $f : \{0,1\}^n \to \{0,1\}$ be ϵ-far from $(n-k)$-symmetric. Then testing f-isomorphism requires at least $\Omega(\log \log k)$ queries.*

In fact, we believe that more is true—perhaps even $\Omega(k)$ queries are required. But the weaker bound (or, indeed, any function that grows with k) is sufficient to complete the qualitative characterization of functions that are isomorphism-testable with a constant number of queries.

The known hardness results on isomorphism testing are all consistent with Conjecture 1. In particular, by the result in [4], we know that testing f-isomorphism requires at least $\Omega(k)$ queries for *almost all* functions f that are ϵ-far from $(n-k)$-symmetric. A simple extension of the proof in [10] shows that for every $(n-k)$-symmetric function f that is ϵ-far from $(n-k+1)$-symmetric, testing f-isomorphism requires $\Omega(\log \log k)$ queries (assuming k/n is bounded away from 1).

Lastly, let us consider another natural definition of partial symmetry that encompasses both symmetric functions and juntas. The function $f : \{0,1\}^n \to \{0,1\}$ is *k-part symmetric* if there is a partition $\mathcal{I} = \{I_1, \ldots, I_k\}$ of $[n]$ such that f is invariant under any permutation π of $[n]$ where $\pi(I_i) = I_i$ for every $i = 1, \ldots, k$. One may be tempted to guess that k-part symmetric functions are efficiently isomorphism-testable. That is not the case, even when $k = 2$. To see this, consider the function $f(x) = x_1 \oplus x_2 \oplus \cdots \oplus x_{n/2}$. This function is 2-part symmetric, but testing isomorphism to f requires $\Omega(n)$ queries [8].

Acknowledgments

We thank Noga Alon, Per Austrin, Irit Dinur, Ehud Friedgut, and Ryan O'Donnell for useful discussions and valuable feedback.

References

[1] José A. Adell and Pedro Jodrá. Exact Kolmogorov and total variation distances between some familiar discrete distributions. *Journal of Inequalities and Applications*, 2006.

[2] Rudolf Ahlswede and Levon H. Khachatrian. The complete intersection theorem for systems of finite sets. *European Journal of Combinatorics*, 18:125–136, 1997.

[3] Noga Alon and Eric Blais. Testing boolean function isomorphism. *Proc. 14th International Workshop on Randomization and Approximation Techniques in Computer Science*, pages 394–405, 2010.

[4] Noga Alon, Eric Blais, Sourav Chakraborty, David García-Soriano, and Arie Matsliah. Nearly tight bounds for testing function isomorphism, 2011. manuscript.

[5] Noga Alon, Eldar Fischer, Ilan Newman, and Asaf Shapira. A combinatorial characterization of the testable graph properties: It's all about regularity. *SIAM Journal on Computing*, 39:143–167, 2009.

[6] Arnab Bhattacharyya, Elena Grigorescu, and Asaf Shapira. A unified framework for testing linear-invariant properties. In *Proc. 51st Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 478–487, 2010.

[7] Eric Blais. Testing juntas nearly optimally. In *Proc. 41st Annual ACM Symposium on Theory of Computing (STOC)*, pages 151–158, 2009.

[8] Eric Blais, Joshua Brody, and Kevin Matulef. Property testing lower bounds via communication complexity. In *Proc. 26th Annual IEEE Conference on Computational Complexity (CCC)*, pages 210–220, 2011.

[9] Eric Blais and Daniel Kane. Testing linear functions, 2011. manuscript.

[10] Eric Blais and Ryan O'Donnell. Lower bounds for testing function isomorphism. In *Proc. 25th Conference on Computational Complexity (CCC)*, pages 235–246, 2010.

[11] Sourav Chakraborty, David García-Soriano, and Arie Matsliah. Efficient sample extractors for juntas with applications. *Automata, Languages and Programming*, pages 545–556, 2011.

[12] Sourav Chakraborty, David García-Soriano, and Arie Matsliah. Nearly tight bounds for testing function isomorphism. In *Proc. 22nd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1683–1702, 2011.

[13] S.R. Das and C.L. Sheng. On detecting total or partial symmetry of switching functions. *IEEE Trans. on Computers*, C-20(3):352–355, 1971.

[14] I. Diakonikolas, H.K. Lee, K. Matulef, K. Onak, R. Rubinfeld, R.A. Servedio, and A. Wan. Testing for concise representations. In *Proc. 48th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 549–558, 2007.

[15] Irit Dinur and Shmuel Safra. On the hardness of approximating minimum vertex cover. *Annals of Mathematics*, 162(1):439–485, 2005.

[16] Paul Erdős, Chao Ko, and Richard Rado. Intersection theorems for systems of finite sets. *The Quarterly Journal of Mathematics*, 12(1):313–320, 1961.

[17] Eldar Fischer, Guy Kindler, Dana Ron, Shmuel Safra, and Alex Samorodnitsky. Testing juntas. *Journal of Computer and System Sciences*, 68(4):753–787, 2004.

[18] Peter Frankl. The Erdős-Ko-Rado theorem is true for $n = ckt$. In *Combinatorics (Proc. Fifth Hungarian Colloquium, Keszthely)*, volume 1, pages 365–375, 1976.

[19] Ehud Friedgut. On the measure of intersecting families, uniqueness and stability. *Combinatorica*, 28(5):503–528, 2008.

[20] Oded Goldreich. On testing computability by small width OBDDs. *Proc. 14th International Workshop on Randomization and Approximation Techniques in Computer Science*, pages 574–587, 2010.

[21] Oded Goldreich, editor. *Property Testing: Current Research and Surveys*, volume 6390 of *LNCS*. Springer, 2010.

[22] Oded Goldreich, Shafi Goldwasser, and Dana Ron. Property testing and its connection to learning and approximation. *Journal of the ACM*, 45(4):653–750, 1998.

[23] Tali Kaufman and Madhu Sudan. Algebraic property testing: the role of invariance. In *Proc. 40th Annual ACM Symposium on Theory of Computing (STOC)*, pages 403–412, 2008.

[24] Christoph Meinel and Thorsten Theobald. *Algorithms and Data Structures in VLSI Design.* Springer, 1998.

[25] Dana Ron. Algorithmic and analysis techniques in property testing. *Foundations and Trends in Theoretical Computer Science*, 5:73–205, 2010.

[26] Ronitt Rubinfeld and Asaf Shapira. Sublinear time algorithms. *Electronic Colloquium on Computational Complexity (ECCC)*, 18, 2011. TR11-013.

[27] Ronitt Rubinfeld and Madhu Sudan. Robust characterizations of polynomials with applications to program testing. *SIAM Journal on Computing*, 25(2):252–271, 1996.

[28] Claude E. Shannon. The synthesis of two-terminal switching circuits. *Bell System Technical Journal*, 28(1):59–98, 1949.

[29] Spario Y. T. Soon. Binomial approximation for dependent indicators. *Statistica Sinica*, 6:703–714, 1996.

[30] Richard M. Wilson. The exact bound in the Erdős-Ko-Rado theorem. *Combinatorica*, 4(2–3):247–257, 1984.

A Properties of symmetric influence

A.1 Fourier representation of symmetric influence

For convenience, we consider functions whose ranges are $\{-1,1\}$ instead of $\{0,1\}$. Then, the symmetric influence of a function can be expressed as follows.

Proposition 2. *Given a Boolean function $f : \{0,1\}^n \to \{-1,1\}$ and a set $J \subseteq [n]$, the symmetric influence of J with respect to f can also be computed as*

$$\mathrm{SymInf}_f(J) = \tfrac{1}{2} \sum_{S \subseteq [n]} \mathop{\mathbf{Var}}_{\pi \in \mathcal{S}_J}[\widehat{f}(\pi S)]$$

where $\widehat{f}(S)$ is the Fourier coefficient of f for the set $S \subseteq [n]$, and $\pi S = \{\pi(i) \mid i \in S\}$.

The proposition indicates that the symmetric influence of any set J can be computed as a function of the variance of the Fourier coefficients of the function in the different layers. Each layer here refer to all the Fourier coefficients of sets which share the intersection with $[n] \setminus J$ and the intersection size with J, resulting in $(|J|+1)2^{n-|J|}$ different layers.

The key to proving this proposition is the following basic result on linear functions. Recall that for a set $S \subseteq [n]$, the function $\chi_S : \{0,1\}^n \to \{-1,1\}$ is defined by $\chi_S(x) = (-1)^{\sum_{i \in S} x_i}$.

Lemma 8. *Fix $J, S, T \subseteq [n]$. Then*

$$\mathop{\mathbf{E}}_{x \in \{0,1\}^n, \pi \in \mathcal{S}_J}[\chi_S(x) \cdot \chi_T(\pi x)] = \begin{cases} \binom{|J|}{|S \cap J|}^{-1} & \text{if } \exists \pi \in \mathcal{S}_J,\ \pi S = T \\ 0 & \text{otherwise.} \end{cases}$$

Proof. For any vector $x \in \{0,1\}^n$, any set $S \subseteq [n]$, and any permutation $\pi \in \mathcal{S}_n$, we have the identity $\chi_S(\pi x) = \chi_{\pi^{-1}S}(x)$. So

$$\mathop{\mathbf{E}}_{x \in \{0,1\}^n, \pi \in \mathcal{S}_J}[\chi_S(x) \cdot \chi_T(\pi x)] = \mathop{\mathbf{E}}_{x,\pi}[\chi_S(x)\chi_{\pi^{-1}T}(x)] = \mathop{\mathbf{E}}_{\pi}\left[\mathop{\mathbf{E}}_{x}[\chi_S(x)\chi_{\pi^{-1}T}(x)]\right] .$$

But $\mathbf{E}_x[\chi_S(x)\chi_{\pi^{-1}T}(x)] = \mathbf{1}[S = \pi^{-1}T]$, so we also have

$$\mathop{\mathbf{E}}_{x \in \{0,1\}^n, \pi \in \mathcal{S}_J}[\chi_S(x) \cdot \chi_T(\pi x)] = \Pr_{\pi \in \mathcal{S}_J}[S = \pi^{-1}T] = \Pr_{\pi \in \mathcal{S}_J}[\pi S = T] .$$

The identity $\pi S = T$ holds iff the permutation π satisfies $\pi(i) \in T$ for every $i \in S$. Since we only permute elements from J, the sets S and T must agree on the elements of $[n] \setminus J$. If this is not the case, or if the intersection of the sets with J is not of the same size, no such permutation exists. Otherwise, this event occurs if the elements of $S \cap J$ are mapped to the exact locations of $T \cap J$. This holds for one out of the $\binom{|J|}{|S \cap J|}$ possible sets of locations, each with equal probability. □

Proof of Proposition 2. By appealing to the fact that f is $\{-1,1\}$-valued, we have that

$$\Pr_{x,\pi}[f(x) \neq f(\pi x)] = \frac{1}{4} \mathop{\mathbf{E}}_{x,\pi}[f(x)^2 + f(\pi x)^2 - 2f(x)f(\pi x)].$$

Applying linearity of expectation and Parseval's identity, we obtain

$$\mathop{\mathbf{E}}_{x,\pi}[f(x)^2 + f(\pi x)^2 - 2f(x)f(\pi x)] = 2 \sum_{S \subseteq [n]} \hat{f}(S)^2 - 2 \sum_{S,T \subseteq [n]} \hat{f}(S)\hat{f}(T) \mathop{\mathbf{E}}_{x,\pi}[\chi_S(x)\chi_T(\pi x)] .$$

Fix any $S \subseteq [n]$. By Lemma 8,

$$\sum_{T \subseteq [n]} \hat{f}(T) \mathop{\mathbf{E}}_{x,\pi}[\chi_S(x)\chi_T(\pi x)] = \sum_{\pi \in \mathcal{S}_J} \frac{\hat{f}(\pi S)}{\binom{|J|}{|S \cap J|}} = \mathop{\mathbf{E}}_{\pi \in \mathcal{S}_J}[\hat{f}(\pi S)] \ .$$

Given this equality,

$$\sum_{S,T \subseteq [n]} \hat{f}(S)\hat{f}(T) \mathop{\mathbf{E}}_{x,\pi}[\chi_S(x)\chi_T(\pi x)] = \sum_S \hat{f}(S) \mathop{\mathbf{E}}_{\pi \in \mathcal{S}_J}[\hat{f}(\pi S)] \ .$$

By applying some elementary manipulation, we now get

$$\begin{aligned} \Pr_{x,\pi}[f(x) \neq f(\pi x)] &= \frac{1}{2}\sum_S \hat{f}(S)\left(\hat{f}(S) - \mathop{\mathbf{E}}_{\pi}[\hat{f}(\pi S)]\right) = \\ &= \frac{1}{2}\sum_S (\mathop{\mathbf{E}}_{\pi}[\hat{f}(\pi S)^2] - \mathop{\mathbf{E}}_{\pi}[\hat{f}(\pi S)]^2) = \frac{1}{2}\sum_S \mathop{\mathbf{Var}}_{\pi}[\hat{f}(\pi S)] \ . \end{aligned}$$

□

A.2 Monotonicity of symmetric influence

Lemma 4 (Restated)**.** *For any function $f : \{0,1\}^n \to \{0,1\}$ and any sets $J \subseteq K \subseteq [n]$,*

$$\mathrm{SymInf}_f(J) \leq \mathrm{SymInf}_f(K) \ .$$

Proof. Fix a function f and two sets $J, K \subseteq [n]$ so that $J \subseteq K$. We have seen before that the symmetric influence can be computed in layers, where each layer is determined by the Hamming weight and the elements outside the set we are considering. Using the fact that $\mathbf{Var}(X) = \Pr[X = 0] \cdot \Pr[X = 1]$, the symmetric influence is twice the expected variance over all the layers (considering also the size of the layers). Using the same notation as before,

$$\begin{aligned} \mathrm{SymInf}_f(J) &= \frac{1}{2^n}\sum_z \sum_w |L^w_{\overline{J} \leftarrow z}| \cdot 2 \mathop{\mathbf{Var}}_x[f(x) \mid x \in L^w_{\overline{J} \leftarrow z}] \\ &= 2 \cdot \mathop{\mathbf{E}}_y\left[\mathop{\mathbf{Var}}_x[f(x) \mid x \in L^{|y|}_{\overline{J} \leftarrow y_{\overline{J}}}]\right] \ . \end{aligned}$$

A key observation is that since $\overline{K} \subseteq \overline{J}$, the layers determined when considering J are a refinement of the layers determined when considering K. Together with the fact that $\mathbf{Var}(X) = \Pr[X = 0] \cdot \Pr[X = 1]$ is a concave function in the range $[0,1]$, we can apply Jensen's inequality on each layer before and after the refinement to get the desired inequality. More precisely, for every $z \in \{0,1\}^{|\overline{K}|}$ and $0 \leq w \leq n$,

$$\mathop{\mathbf{Var}}_x[f(x) \mid x \in L^w_{\overline{K} \leftarrow z}] \geq \mathop{\mathbf{E}}_y\left[\mathop{\mathbf{Var}}_x[f(x) \mid x \in L^w_{\overline{J} \leftarrow y_{\overline{J}}}] \mid y \in L^w_{\overline{K} \leftarrow z}\right] \ .$$

Averaging this over all layers, we get the desired result. □

A.3 Weak sub-additivity of symmetric influence

In this section we prove that symmetric influence satisfies weak sub-additivity. It might be tempting to think that strong sub-additivity holds, as in the standard notion of influence, however this is not the case. For example, consider the function $f(x) = f_1(x_J) \oplus f_2(x_K)$ for some partition $[n] = J \cup K$ and two randomly chosen symmetric functions f_1, f_2. Since f is far from symmetric, $\mathrm{SymInf}_f([n]) = \mathrm{SymInf}_f(J \cup K) > 0$ while $\mathrm{SymInf}_f(J) = \mathrm{SymInf}_f(K) = 0$.

The additive factor of $c\sqrt{\gamma}$ in Lemma 5 is derived from the distance between the two distributions $\pi_{J\cup K}x$ and $\pi_J\pi_K x$, for a random $x \in \{0,1\}^n$ and random permutations from $\mathcal{S}_{J\cup K}, \mathcal{S}_J, \mathcal{S}_K$. When the sets J and K are large, the distance between these distributions is relatively small which therefore result in this weak sub-additivity property.

The analysis of the lemma is done using hypergeometric distributions, and the distance between them. Let $\mathcal{H}_{n,m,k}$ be the hypergeometric distribution obtained when we pick k balls out of n, m of which are red, and count the number of red balls we obtained. Let $\mathrm{d_{TV}}(\cdot,\cdot)$ denote the statistical distance between two distributions. The following two lemmas would be useful for our proof.

Lemma 9. *Let $J, K \subseteq [n]$ be two sets and π, π_J, π_K be permutations chosen uniformly at random from $\mathcal{S}_{J\cup K}, \mathcal{S}_J, \mathcal{S}_K$, respectively. For a fixed $x \in \{0,1\}^n$, we define $\mathcal{D}_{\pi x}$ and $D_{\pi_J\pi_K x}$ as the distribution of πx and $\pi_J\pi_K x$, respectively. Then,*

$$\mathrm{d_{TV}}(D_{\pi x}, D_{\pi_J\pi_K x}) = \mathrm{d_{TV}}(\mathcal{H}_{|J\cup K|,|x_{J\cup K}|,|K\setminus J|}, \mathcal{H}_{|K|,|x_K|,|K\setminus J|})$$

holds.

Lemma 10. *Let n, m, n', m', k be non-negative integers with $k, n' \le \gamma n$ for some $\gamma \le \frac{1}{2}$. Suppose that $|m - \frac{n}{2}| \le t\sqrt{n}$ and $|m' - \frac{n'}{2}| \le t\sqrt{n'}$ hold for some $t \le \frac{1}{100\sqrt{\gamma}}$. Then,*

$$\mathrm{d_{TV}}(\mathcal{H}_{n,m,k}, \mathcal{H}_{n-n',m-m',k}) \le c_{10}(1+t)\gamma \ .$$

holds for some universal constant c_{10}.

We first show how these lemmas imply the proof of Lemma 5, and will afterwards prove them.

Lemma 5 (Restated)**.** *There is a universal constant c such that, for any constant $0 < \gamma < 1$, a function $f : \{0,1\}^n \to \{0,1\}$ and sets $J, K \subseteq [n]$ of size at least $(1-\gamma)n$,*

$$\mathrm{SymInf}_f(J \cup K) \le \mathrm{SymInf}_f(J) + \mathrm{SymInf}_f(K) + c\sqrt{\gamma} \ .$$

Proof. Let π, π_J and π_K be as in Lemma 9 and fix $x \in \{0,1\}^n$ to be some input.

$$\begin{aligned}\Pr_{\pi}[f(x) \ne f(\pi x)] &\le \Pr_{\pi_J,\pi_K}[f(x) \ne f(\pi_J\pi_K x)] + \mathrm{d_{TV}}(\mathcal{D}_{\pi x}, \mathcal{D}_{\pi_J\pi_K x}) \\ &\le \Pr_{\pi_K}[f(x) \ne f(\pi_K x)] + \Pr_{\pi_J,\pi_K}[f(\pi_K x) \ne f(\pi_J\pi_K x)] + \mathrm{d_{TV}}(\mathcal{D}_{\pi x}, \mathcal{D}_{\pi_J\pi_K x})\end{aligned}$$

By summing over all possible inputs x we have

$$\begin{aligned}\mathrm{SymInf}_f(J \cup K) &= \Pr_{x,\pi}[f(x) \ne f(\pi x)] = \frac{1}{2^n}\sum_x \Pr_{\pi}[f(x) \ne f(\pi x)] \\ &\le \mathrm{SymInf}_f(J) + \mathrm{SymInf}_f(K) + \frac{1}{2^n}\sum_x \mathrm{d_{TV}}(\mathcal{D}_{\pi x}, \mathcal{D}_{\pi_J\pi_K x}) \ .\end{aligned}$$

By applying Lemma 9 over each input x, it suffices to show that

$$\frac{1}{2^n}\sum_x \mathrm{d_{TV}}(\mathcal{D}_{\pi x}, \mathcal{D}_{\pi_J \pi_K x}) = \frac{1}{2^n}\sum_x \mathrm{d_{TV}}(\mathcal{H}_{|J\cup K|,|x_{J\cup K}|,|K\setminus J|}, \mathcal{H}_{|K|,|x_K|,|K\setminus J|}) \le c\sqrt{\gamma}\ . \tag{1}$$

Ideally, we would like to apply Lemma 10 on every input x and get the desired result, however this is not possible as some inputs does not satisfy the requirements of the lemma. Therefore, we perform a slightly more careful analysis. Let us choose $c \ge 2$ and assume $\gamma \le \frac{1}{4}$ (as otherwise the claim trivially holds). Fix $\gamma' = \gamma/(1-\gamma) \le \frac{1}{2}$ and $t = \frac{1}{100\sqrt{\gamma'}}$. We first note that regardless of x, the required conditions on the size of the sets hold. To be exact, $|J \setminus K| \le \gamma'|J \cup K|$ and $|K \setminus J| \le \gamma'|J \cup K|$ since $|J \cup K| \ge (1-\gamma)n$ and $|J \setminus K| \le |\overline{K}| \le \gamma n$ (and similarly $|K \setminus J| \le \gamma n$).

We say an input x is *good* if it satisfies the other conditions of Lemma 10. That is, both $\left||x_{J\cup K}| - \frac{|J\cup K|}{2}\right| \le t\sqrt{|J \cup K|}$ and $\left||x_{J\setminus K}| - \frac{|J\setminus K|}{2}\right| \le t\sqrt{|J \setminus K|}$ hold. Otherwise we call such x *bad*. From the Chernoff bound and the union bound, the probability that x is bad is at most $4\exp(-2t^2) \le 4\exp\left(-\frac{1}{5000\gamma'}\right) \le c'\gamma$ for some constant c' (notice that $\gamma' \le 2\gamma$).

By applying Lemma 10 over the good inputs we get

$$(1) \le \frac{1}{2^n}\sum_{x:bad} 1 + \frac{1}{2^n}\sum_{x:good} c_{10}(1+t)\gamma \le c'\gamma + c_{10}(1+t)\gamma \le c\sqrt{\gamma}$$

for some constant c, as required. □

Proof of Lemma 9. Since both distributions $D_{\pi x}$ and $D_{\pi_J\pi_K x}$ only modify coordinates in $J \cup K$, we can ignore all other coordinates. Moreover, it is in fact suffices to look only at the number of ones in the coordinates of $K \setminus J$ and $J \cup K$, which completely determines the distributions. Let D_z denote the uniform distribution over all elements $y \in \{0,1\}^n$ such that $|y| = |x|$, $y_{\overline{J\cup K}} = x_{\overline{J\cup K}}$ and $|y_{K\setminus J}| = z$ (which also fixes the number of ones in y_J). Notice that this is well defined only for values of z such that $\max\{0, |x_{J\cup K}| - |J|\} \le z \le \min\{|x_{J\cup K}|, |K \setminus J|\}$.

Given this notation, $D_{\pi x}$ can be looked at as choosing $z \sim \mathcal{H}_{|J\cup K|,|x_{J\cup K}|,|K\setminus J|}$ and returning $y \sim D_z$. This is because we apply a random permutation over all elements of $J \cup K$, and therefore the number of ones inside $K \setminus J$ is indeed distributed like z. Moreover, the order inside both sets $K \setminus J$ and J is uniform.

The distribution $D_{\pi_J\pi_K x}$ can be looked at as choosing $z \sim \mathcal{H}_{|K|,|x_K|,|K\setminus J|}$ and returning $y \sim D_z$. The number of ones in $K \setminus J$ is determined already after applying π_K. It is distributed like z as we care about the choice of $|K \setminus J|$ out of the $|K|$ elements, and $|x_K|$ of them are ones (and their order is uniform). Later, we apply a random permutation π_J over all other relevant coordinates, so the order of elements in J is also uniform.

Since the distributions D_z are disjoint for different values of z, this implies that the distance between the two distributions $D_{\pi x}$ and $D_{\pi_J\pi_K x}$ depends only on the number of ones chosen to be inside $K \setminus J$. Therefore we have

$$\mathrm{d_{TV}}(D_{\pi x}, D_{\pi_J\pi_K x}) = \mathrm{d_{TV}}(\mathcal{H}_{|J\cup K|,|x_{J\cup K}|,|K\setminus J|}, \mathcal{H}_{|K|,|x_K|,|K\setminus J|})$$

as required. □

Proof of Lemma 10. Our proof uses the connection between hypergeometric distribution and the binomial distribution, which we denote by $\mathcal{B}_{n,p}$ (for n experiments, each with success probability p). By the triangle inequality we know that

$$\mathrm{d_{TV}}(\mathcal{H}_{n,m,k}, \mathcal{H}_{n-n',m-m',k}) \le \mathrm{d_{TV}}(\mathcal{H}_{n,m,k}, \mathcal{B}_{k,p}) + \mathrm{d_{TV}}(\mathcal{B}_{k,p}, \mathcal{B}_{k,p'}) + \mathrm{d_{TV}}(\mathcal{B}_{k,p'}, \mathcal{H}_{n-n',m-m',k}) \tag{2}$$

where $p = \frac{m}{n}$ and $p' = \frac{m-m'}{n-n'}$. In order to bound the distances we just introduced, we use the following two lemmas.

Lemma 11 (Example 1 in [29])**.** $\mathrm{d}_{\mathrm{TV}}(\mathcal{H}_{n,m,k}, \mathcal{B}_{k,p}) \leq \frac{k}{n}$ *holds for* $p = \frac{m}{n}$.

Lemma 12 ([1])**.** *Let* $0 < p < 1$ *and* $0 < \delta < 1 - p$*. Then,*

$$\mathrm{d}_{\mathrm{TV}}(\mathcal{B}_{n,p}, \mathcal{B}_{n,p+\delta}) \leq \frac{\sqrt{e}}{2} \frac{\tau_{n,p}(\delta)}{(1 - \tau_{n,p}(\delta))^2}$$

provided $\tau_{n,p}(\delta) = \delta\sqrt{\frac{n+2}{2p(1-p)}} < 1$.

Before using the above lemmas, we analyze some of the parameters. First, when $k = 0$ the lemma trivially holds and we therefore assume $k \geq 1$. Notice that this implies that $n\gamma \geq k \geq 1$. The probability p is known to be relatively close to half. To be exact, $|p - \frac{1}{2}| \leq t\sqrt{n}/n \leq \frac{1}{100\sqrt{n\gamma}} \leq \frac{1}{100}$ and therefore $\frac{1}{p(1-p)} < 6$. Assume $p \leq p'$ and let $\delta = p' - p$ (the other case can be treated in the same manner). We first bound δ as follows.

$$\begin{aligned}
\delta &= \frac{mn' - nm'}{n(n-n')} \leq \frac{1}{n(n-n')}\left(\left(\frac{n}{2} + t\sqrt{n}\right)n' - n\left(\frac{n'}{2} - t\sqrt{n'}\right)\right) \\
&= \frac{t(n\sqrt{n'} + \sqrt{n}n')}{n(n-n')} \leq \frac{2t\sqrt{\gamma}n^{3/2}}{(1-\gamma)n^2} \leq 4t\sqrt{\frac{\gamma}{n}} \quad (\text{from } \gamma \leq \frac{1}{2}) .
\end{aligned}$$

Then, $\tau_{k,p}(\delta)$ in Lemma 12 can be bounded by

$$\begin{aligned}
\tau_{k,p}(\delta) &\leq 4t\sqrt{\frac{\gamma}{n}}\sqrt{\frac{k+2}{2p(1-p)}} \leq 4t\sqrt{\frac{3\gamma(k+2)}{n}} \quad (\text{from } \tfrac{1}{p(1-p)} < 6) \\
&\leq 12t\sqrt{\gamma k/n} \leq 12t\gamma \quad (\text{from } 1 \leq k \leq \gamma n) .
\end{aligned}$$

Note that, from the assumption, we have $\tau_{k,p}(\delta) \leq \frac{1}{2}$. By Lemmas 11 and 12, we have

$$\begin{aligned}
(2) &\leq \frac{k}{n} + \frac{\sqrt{e}}{2}\frac{\tau_{k,p}(\delta)}{(1-\tau_{k,p}(\delta))^2} + \frac{k}{n-n'} \\
&\leq 3\gamma + 2\sqrt{e} \cdot 12t\gamma \quad (\text{from } \tau_{k,p}(\delta) \leq \frac{1}{2}) \\
&\leq c_{10}(1+t)\gamma
\end{aligned}$$

for some universal constant c_{10}. □

B Testing partial symmetry

B.1 Analysis of Find-Asymmetric-Set

In this section we prove there exists an algorithm Find-Asymmetric-Set, which satisfies Lemma 6.

Suppose that we have two inputs $x, y \in \{0,1\}^n$ with $x_J = y_J, |x| = |y|$ such that $f(x) \neq f(y)$. Given such inputs, we know there exists some asymmetric variable outside of J. In order to efficiently find a set from a partition $\mathcal{I}$ which contains such a variable, we will use binary search over the sets. First, we construct a refinement $\mathcal{J}$ of $\mathcal{I}$. Every set of $\mathcal{I} \setminus \{W\}$ is partitioned further into parts so that each part has size at most $\lceil |W|/4 \rceil$. Let $t = |\mathcal{J} \setminus \{W\}|$ be the number of parts in $\mathcal{J}$ excluding the workspace. Notice that the number of parts is at most $t \leq r + 4n/|W| = O(r)$. Then, we construct a series of inputs $x^0 = x, x^1, \ldots, x^t = y$ by each step permuting only elements from some set $I \in \mathcal{J} \setminus \{W\}$ and the workspace W (that is, applying a permutation from $\mathcal{S}_{I \cup W}$). In each such step, we guarantee that $x^i_I = y_I$ for one more set $I \in \mathcal{J} \setminus \{W\}$, and therefore after (at most) t steps we would reach y (notice that we can choose the last step such that $x^t_W = y_W$ as the Hamming weight of all the inputs in the sequence is identical).

Using this construction, we can now describe the algorithm Find-Asymmetric-Set as follows.

Algorithm 4 Find-Asymmetric-Set$(f, \mathcal{I}, J, W)$

```
Generate $x \in \{0,1\}^n$ and $\pi \in \mathcal{S}_{\overline{J}}$ uniformly at random.
if $f(x) \neq f(\pi x)$ then
    Define $x^0, \ldots, x^t$.
    Perform binary search on $x = x^0, \ldots, x^t = y$, and find $i$ such that $f(x^{i-1}) \neq f(x^i)$.
    return the only part $I \in \mathcal{I} \setminus \{W\}$ such that $x^{i-1}_I \neq x^i_I$.
return $\emptyset$.
```

Proof of Lemma 6. Since we perform binary search over the sequence $x^0, \ldots, x^t$, the query complexity of the algorithm is indeed $O(\log t) = O(\log r)$. Also, it is easy to verify that we only output an empty set or a part in $\mathcal{I} \setminus \{W\}$ disjoint to J (as $x_J = y_J$).

Two random inputs x and $y := \pi x$, for $\pi \in \mathcal{S}_J$, satisfy $f(x) \neq f(y)$ with probability $\mathrm{SymInf}_f(\overline{J})$. Thus, it suffices to show that we can always define a sequence of $x^0, \ldots, x^t$, given that $|W| \geq \frac{n}{2r}$. In order to see this is always feasible, we consider the sequence after already defining $x^0, \ldots, x^i$, showing we can define x^{i+1}.

Let $\mathcal{J}^+ = \{I \in \mathcal{J} \mid |x^i_I| > |y_I|\}$ and $\mathcal{J}^- = \{I \in \mathcal{J} \mid |x^i_I| < |y_I|\}$ denote the sets which require increasing or decreasing the Hamming weight of x_W respectively, when applying a permutation from $\mathcal{S}_{I \cup W}$ to ensure $x^{i+1}_I = y_I$. Notice that we ignore sets I for which $|x^i_I| = |y_I|$, as they do not impact the Hamming weight of x^i_W. If $|\mathcal{J}^+| > 0$ and $|\mathcal{J}^-| > 0$, then since $\max(|x^i_W|, |W| - |x^i_W|) \geq \lceil |W|/2 \rceil$ and the size of every set $I \in \mathcal{J} \setminus \{W\}$ is at most $\lceil |W|/4 \rceil$, there must exists a set we can use to define x^{i+1}. On the other hand, if $|\mathcal{J}^+| = 0$ for example, then we can define x^{i+1} using any set from $\mathcal{J}^-$ as $|x^i_W| - |y_W| = -\sum_{I \in \mathcal{J} \setminus \{W\}} |x^i_I| - |y_I|$ (recall that $|x| = |x^i| = |y|$).

It remains to show that when W contains no asymmetric variables and we output a part $I \in \mathcal{I} \setminus \{W\}$, I contains an asymmetric variable. Suppose that the output I is the part which was modified between x^{i-1} and x^i. Then, since $f(x^{i-1}) \neq f(x^i), |x^{i-1}| = |x^i|$, and x^{i-1} and x^i differ only on $I \cup W$, an asymmetric variable exists in $I \cup W$ and we know it is not in W. □

B.2 Proof of Lemma 7

We first note that when the number of parts r is bigger then n, we simply partition into the n single-element sets and the lemma trivially holds. For $0 \leq t \leq 1$, let $\mathcal{F}_t = \{J \subseteq [n] : \mathrm{SymInf}_f(\overline{J}) < t\epsilon,\ |J| \leq 5kn/r\}$ be the family of all sets which are not too big and whose complement has symmetric influence of at most $t\epsilon$. (Notice that with high probability, the union of any k sets in the partition would have size smaller than $5kn/r$, and therefore we assume this is the case from this point on.) Our first observation is that for small enough values of t, $\mathcal{F}_t$ is a $(k+1)$-intersecting family. Indeed, for any sets $J, K \in \mathcal{F}_{1/3}$,

$$\mathrm{SymInf}_f(\overline{J \cap K}) = \mathrm{SymInf}_f(\overline{J} \cup \overline{K}) \leq \mathrm{SymInf}_f(\overline{J}) + \mathrm{SymInf}_f(\overline{K}) + c\sqrt{5k/r} < 2\epsilon/3 + \epsilon/9 < \epsilon\ .$$

Since f is ϵ-far from $(n-k)$-symmetric, every set $S \subseteq [n]$ of size $|S| \le k$ satisfies $\mathrm{SymInf}_f(\overline{S}) \ge \epsilon$. So $|J \cap K| > k$.

We consider two cases separately: when $\mathcal{F}_{1/3}$ contains a set of size less than $2k$; and when it does not. The first case is identical to the proof of Lemma 2 and hence we do not elaborate on it.

In the second case, which also resembles the proof of Lemma 2, we claim that $\mathcal{F}_{1/9}$ is a $2k$-intersecting family. If this was not the case, we could find sets $J, K \in \mathcal{F}_{1/9}$ such that $|J \cap K| < 2k$ and $\mathrm{SymInf}_f(\overline{J \cap K}) \le \mathrm{SymInf}_f(\overline{J}) + \mathrm{SymInf}_f(\overline{K}) + \epsilon/9 < \epsilon/3$, contradicting our assumption.

Let $J \subseteq [n]$ be the union of k parts in $\mathcal{I}$. Since $\mathcal{I}$ is a random partition, J is a random subset obtained by including each element of $[n]$ in J independently with probability $p = k/r < \frac{1}{2k+1}$. To bound the probability that J contains some element from $\mathcal{F}_{1/9}$, we define $\mathcal{F}'_{1/9}$ to be all the sets that contain a member from $\mathcal{F}_{1/9}$. Since $\mathcal{F}'_{1/9}$ is also a $2k$-intersecting family, by Theorem 4, for every such J of size at most $5kn/r$, $\Pr[\mathrm{SymInf}_f(\overline{J}) < \frac{\epsilon}{9}] = \Pr[J \in \mathcal{F}_{1/9}] \le \mu_{k/r}(\mathcal{F}'_{1/9}) \le (k/r)^{2k}$. Applying the union bound over all possible choices for k parts, f will not satisfy the condition of the lemma with probability at most $\binom{r}{k}\left(\frac{k}{r}\right)^{2k} = O(k^{-k})$, which completes the proof of the lemma.

C Isomorphism testing and sampling partially symmetric functions

C.1 Properties of the sampling distribution

We start this section with the following observation. When the number of parts r reaches n (or alternatively when $k = \Omega(\sqrt{n})$), we consider the partition of $[n]$ into the n single-element sets. Notice that when this is the partition, then in fact $\mathcal{D}_{\mathcal{I}}^{W}$ is identical to $\mathcal{D}^{*}_{k,n}$, making the following proposition trivial. Therefore, in the proof we assume that $r < n$ and $k = O(\sqrt{n})$.

Proof of Proposition 1. We start with the first part of the proposition, showing x is almost uniform. Consider the following procedure to generate a random $\mathcal{I}, W$ and x. We draw a random Hamming weight $w \sim \mathcal{B}_{n,1/2}$ and define x' to be the input consisting of w ones followed by $n-w$ zeros. We choose a random partition $\mathcal{I}'$ of $[n]$ into r *consecutive* parts $I_1, \ldots, I_r$ (i.e., $I_1 = \{1, 2, \ldots, |I_1|\}$ and $I_r = \{n - |I_r| + 1, \ldots, n\}$) according to the typical distribution of sizes in a random partition. Let the workspace W' be the only part which contains the coordinate w (or I_1 if $w = 0$). We now apply a random permutation over x', $\mathcal{I}'$ and W' to get x, $\mathcal{I}$ and W.

It is clear the above procedure outputs a uniform x as we applied a random permutation over x', which had a binomial Hamming weight. The choice of $\mathcal{I}$ was also done at random, considering the applied permutation over $\mathcal{I}'$. The only difference is then in the choice of the workspace W, which can only be reflected in its size. However, when $r = o(\sqrt{n})$ we will choose the middle part as the workspace with probability $1 - o(1)$, regardless of its size. In the remaining cases, since there are $n/r = \Omega(\sqrt{n})$ parts, the possible parts to be chosen as workspace are a small fraction among all parts, and therefore W would be $o(1)$-close to being a random part.

Proving the second property of the proposition, we also consider two cases. When $r = o(\sqrt{n})$, with probability $1 - o(1)$, the workspace would have size $\omega(\sqrt{n})$ and also $w = n/2 + O(\sqrt{n})$. In such a case, the $r-1$ parts (excluding the workspace) would be half zeros and half ones, and the marginal distribution over the number of ones in J would be $\mathcal{H}_{r-1,(r-1)/2,k}$ (assuming the elements of J are separated by $\mathcal{I}$, which happens with probability $1 - o(1)$). By Lemma 11, the distance between this distribution and $\mathcal{B}_{k,1/2}$ is bounded by $k/r < c/k$ for our choice of $0 < c < 1$. Since there is no restriction on the ordering of the sets, this is also the distance from uniform over $\{0,1\}^k$ as required.

In the remaining case where $r = \Omega(\sqrt{n})$, we can use the same arguments and also apply Lemma 12 with the distributions $\mathcal{B}_{k,1/2}$ and $\mathcal{B}_{k,1/2+\delta}$ for $\delta = O(1/\sqrt{n})$, implying the distance between these two distributions is at most $o(1)$. Combining this with the distance to $\mathcal{H}_{r-1,(r-1)(1/2+\delta),k}$ we get again a total distance of $k/r + o(1) < c/k$ for our choice of $0 < c < 1$. □

C.2 Analysis of Partially-Symmetric-Isomorphism-Test

The analysis of the algorithm is based on the fact that functions which passes the Partially-Symmetric-Test satisfy some conditions, and in particularly are closed to being partially symmetric. We therefore start with the following lemma.

Lemma 13. *Let g be a function ϵ-close to being $(n-k)$-symmetric which passed the* Partially-Symmetric-Test(g,k,ϵ). *In addition, let $\mathcal{I}, W$ and J be the partition, workspace and identified parts used by the algorithm. With probability at least $9/10$, there exists a function h which satisfies the following properties.*

- *h is 4ϵ-close to g, and*
- *h is $(n-k)$-symmetric whose asymmetric variables are contained in J and separated by $\mathcal{I}$.*

Proof. Let g^* be the $(n-k)$-symmetric function closest to g (which can be f itself, up-to some isomorphism) and R be the set of (at most) k asymmetric variables of g^*. By Lemma 3 and our assumption over g,

$$\mathrm{SymInf}_g(\overline{R}) \le 2 \cdot \mathrm{dist}(g, g^*) \le 2\epsilon \ .$$

Notice however that R is not necessarily contained in J and therefore g^* is not a good enough candidate for h. Let $U = R \cap J$ be the intersection of the asymmetric variables of g^* and the sets identified by the algorithm. In order to show that g is also close to being $\overline{U}$-symmetric, we bound $\mathrm{SymInf}_g(\overline{U})$ using Lemma 5 with the sets $\overline{R}$ and $\overline{J}$. Notice that since $|R| \le k$ and $|J| \le 2kn/r \le \epsilon^2 n/c'$ for our choice of c', we can bound the error term (in the notation of Lemma 5) by $c\sqrt{\gamma} \le c\sqrt{\epsilon^2/c'} \le \epsilon$. We therefore have

$$\mathrm{SymInf}_g(\overline{U}) \le \mathrm{SymInf}_g(\overline{R}) + \mathrm{SymInf}_g(\overline{J}) + \epsilon \le 2\epsilon + \epsilon + \epsilon = 4\epsilon$$

where we know $\mathrm{SymInf}_g(\overline{J}) \le \epsilon$ with probability at least $19/20$ as the algorithm did not reject.

By applying Lemma 3 again, we know there exists a $\overline{U}$-symmetric function h, whose distance to g is bounded by $\mathrm{dist}(g,h) \le 4\epsilon$. Moreover, with probability at least $19/20$, all its asymmetric variables are completely separated by the partition $\mathcal{I}$ (and they were all identified as part of J). □

Given Lemma 13, we are now ready to analyze Partially-Symmetric-Isomorphism-Test.

Proof of Theorem 1. Before analyzing the algorithm we just described, we consider the case where $k > n/10$. Since Theorem 2 does not hold for such k's, we apply the basic algorithm of $O(n \log n/\epsilon)$ random queries, which is applicable testing isomorphism of any given function (since there are $n!$ possible isomorphisms, the random queries will rule out all of them with good probability, assuming we should reject). Since $k = \Omega(n)$, the complexity of this algorithm fits the statement of our theorem.

We start by analyzing the query complexity of the algorithm. The step of Partially-Symmetric-Test performs $O(\frac{k}{\epsilon} \log \frac{k}{\epsilon})$ queries, and therefore the majority of the queries are performed at the sampling stage, resulting in $O(k \log k/\epsilon^2)$ queries as required. In order to prove the correctness of the algorithm, we consider the following cases.

- g is ϵ-far from being isomorphic to f and $\epsilon/1000$-far from being $(n-k)$-symmetric.
- g is ϵ-far from being isomorphic to f but $\epsilon/1000$-close to being $(n-k)$-symmetric.
- g is isomorphic to f.

In the first case, with probability at least 9/10, Partially-Symmetric-Test will reject and so will we, as required. We assume from this point on that Partially-Symmetric-Test did not reject, as it will only reject g which is isomorphic to f with probability at most 1/10, and that we are not in the first case. Notice that these cases match the conditions of Lemma 13, and therefore from this point onward we assume there exists an h satisfying the lemma's properties (remembering we applied the algorithm with $\epsilon/1000$).

In order to bound the distance between h and g in our samples, we use Proposition 1, indicating

$$\Pr_{\mathcal{I},W\in\mathcal{I},x\sim\mathcal{D}_{\mathcal{I}}^{W}}[g(x)\neq h(x)] = \mathrm{dist}(g,h)+o(1/n)\ .$$

By Markov's inequality, with probability at least 9/10, the partition $\mathcal{I}$ and the workspace W satisfy

$$\Pr_{x\sim\mathcal{D}_{\mathcal{I}}^{W}}[g(x)\neq h(x)] \leq 10\cdot\mathrm{dist}(g,h)+o(1/n)\leq 10\cdot 4\epsilon/1000+o(1/n)<\epsilon/20\ .$$

By Proposition 1, if we were to sample h according to $\mathcal{D}_{\mathcal{I}}^{W}$, it should be $\epsilon/20$-close to sampling its core (assuming the partition size is large enough). Combined with the distance between g and h in our samples, we expect our samples to be $\epsilon/20+\epsilon/20=\epsilon/10$ close to sampling h's core.

The last part of the proof is showing that there would be an almost consistent isomorphism of f only when g is isomorphic to f. Notice however that we care only for isomorphisms which map the asymmetric variables of f to the k sets of J. Therefore, the number of different isomorphisms we need to consider is $k!$.

Assume we are in the second case and g is ϵ-far from being isomorphic to f. Let f_π be some isomorphism of f. By our assumptions and Lemma 13,

$$\mathrm{dist}(f_\pi,h)\geq\mathrm{dist}(f_\pi,g)-\mathrm{dist}(g,h)\geq\epsilon-\epsilon/250\ .$$

Each sample we perform would be inconsistent with f_π with probability at least $\epsilon-\epsilon/250-\epsilon/10>8\epsilon/9$. By the Chernoff bounds and the union bound, if we would perform $q=O(k\log k/\epsilon^2)$ queries, we would rule out all $k!$ possible isomorphisms with probability at least 9/10 and reject the function as required.

On the other hand, if g is isomorphic to f, then we know there exists with probability at least 9/10 some isomorphism f_π which maps the asymmetric variables of f into the sets of J, such that

$$\mathrm{dist}(f_\pi,h)\leq\mathrm{dist}(f_\pi,g)+\mathrm{dist}(g,h)\leq\epsilon/500+\epsilon/250\ .$$

For this isomorphism, with high probability much more than $(1-\epsilon/2)$-fraction of the queries would be consistent and we would therefore accept g as we should. □

C.3 Efficient sampler for partially symmetric functions

We first provide the algorithm for efficiently generating a δ-sampler for partially symmetric functions. The algorithm perform its preprocessing by calling Partially-Symmetric-Test. Given the output of the algorithm, we query the function once for each call to the sampler, according to $\mathcal{D}_{\mathcal{I}}^{W}$, and return the result.

Algorithm 5 Partially-Symmetric-Sampler(f, k, δ, η)

```
Perform Partially-Symmetric-Test$(f, k, \eta\delta)$.
Let $\mathcal{I}$ and $W \in \mathcal{I}$ be the partition and workspace used by the algorithm.
Let $J$ be the union of $k$ parts in $\mathcal{I} \setminus \{W\}$ that were identified by the algorithm.
Return the following sampler:
    Choose a random $y \sim \mathcal{D}_{\mathcal{I}}^{W}$
    Let $x \in \{0,1\}^k$ be the value assigned to the parts in $J$
    Yield the triplet $(x, |y| - |x|, f(y))$
```

Proof of Theorem 3. The algorithm for generating the sampler is described by Partially-Symmetric-Sampler, which performs $O(\frac{k}{\eta\delta} \log \frac{k}{\eta\delta})$ preprocessing queries to the function. What remains to be proved is that indeed with good probability, the algorithm returns a valid sampler.

Let h be the function defined in the analysis of Theorem 1, which satisfies the conditions of Lemma 13. Recall that its asymmetric variables were separated by $\mathcal{I}$ and appear in J. Following this analysis and that of Partially-Symmetric-Test, one can see that with probability at least $1 - \eta$ we would not reject f when calling Partially-Symmetric-Test. Moreover, the samples would be $\delta/2$-close to sampling the core of h, which is by itself $\delta/2$-close to f. Therefore, overall our samples would be δ-close to sampling the core of f.

The last part in completing the proof of the theorem is showing that we sample the core with distribution δ-close to $\mathcal{D}_{k,n}^{*}$. By Proposition 1, the total variation distance between sampling the core according to $\mathcal{D}_{k,n}^{*}$ and sampling it according to $\mathcal{D}_{\mathcal{I}}^{W}$ is at most c/k for our choice of $0 < c < 1$, which we can choose it to be at most δ. □

Notice that if the function f is not $(n-k)$-symmetric but still very close (say $(k/\eta\delta)^2$-close), applying the same algorithm will provide a good sampler for an $(n-k)$-symmetric function f' close to f. The main reason is that most likely, we will not query any location of the function where it does not agree with f'.